

TITLE OF INVENTION: "DIAGNOSTIC SCANNING MICROSCOPE FOR
INFORMATION-ENRICHED QUALITATIVE
HISTOPATHOLOGY"

INVENTOR(S): Peter H. Bartels

ATTORNEY: Antonio R. Durando
Registration No. 28,409

Durando Birdwell & Janke, P.L.C.
2929 E. Broadway Blvd.
Tucson, AZ 85716
520-881-9442 Telephone
520-881-9448 Fax

ATTORNEY DOCKET NO.: 6121.019

DIAGNOSTIC SCANNING MICROSCOPE FOR INFORMATION-ENRICHED
QUALITATIVE HISTOPATHOLOGY

BACKGROUND OF THE INVENTION

5

Field of the Invention

[0001] This invention is related in general to the field of light
optical microscopy. In particular, it relates to a method and
apparatus for conducting diagnostic testing of biological tissue
10 with a scanning microscope.

Description of the Related Art

[0002] Changes in the cellular structure of tissue are used to
detect pathologic changes, to assess the progress of precancerous
15 conditions, and to detect cancer. A tissue sample removed from a
patient is typically sectioned and fixed to a slide for staining
and microscopic examination by a pathologist. The morphology of
the tissue (the visually perceptible structure and shape of
features in the tissue) is analyzed to provide a qualitative
20 assessment of its condition and to identify the presence of
pathologic changes, such as may indicate progression towards a
malignancy. For many decades, this visual procedure has been the
diagnostic mainstay of pathology.

[0003] With the advent of computers and sophisticated digital imaging equipment, researchers have extended the realm of histopathology through the use of mechanized procedures for diagnostic and quantitative investigation. For example, U.S.

5 Patent No. 6,204,064 describes a method for measuring quantitatively the progression of a lesion toward malignancy by digitizing the images of clinical samples and analyzing nuclear chromatin texture features in the nuclei captured in the images. Numerical values are assigned to these features and compared to a
10 monotonic progression curve previously established using the same criteria on known clinical samples ranging from normal to malignant tissue. Thus, the procedure provides a quantitative assessment of the condition of the tissue as well as a method for testing the efficacy of chemo preventive drugs or therapeutic
15 treatments.

[0004] In such mechanized procedures, histopathologic sections and/or cytologic preparations are imaged with a microscope, and the images are digitized, stored, and analyzed for nuclear-
20 placement patterns (histometry) or for the spatial and statistical distribution patterns of nuclear chromatin (karyometry). Karyometric assessment is always preceded by image segmentation, whereby each nucleus in an image is identified, outlined, isolated and stored as a separate image. For example,

Figs. 1 and 2 illustrate, respectively, the image of a histopathologic section and the enlarged image of a segmented nucleus and its chromatin pattern. As those skilled in the art readily understand, the nuclear chromatin pattern is an artefact of tissue fixation, but its spatial and statistical distributions are highly reproducible measures of the metabolic and functional state of cells. Thus, nuclear chromatin patterns have always been used in pathology to provide diagnostic clues. For example, the state of differentiation of the nucleus and its metabolic function may be reliably assessed based on a finding that nuclear chromatin is finely dispersed, coarsely aggregated, granular, clumped, or displaced toward the nuclear periphery.

[0005] Many chromatin texture features derived from the optical density of the tissue image have been identified as statistically significant for diagnostic purposes. Accordingly, after a sample is imaged and the image is digitized to provide an optical density value for each image pixel, the information is used in conventional manner first to identify and isolate each nucleus within the sample (image segmentation), and then to analyze chromatin patterns within each nucleus. The optical density recorded for each pixel is used to characterize chromatin features with statistical significance as parameters for identifying changes in the condition of the tissue. These

features are then used much as alphabet letters can be used to identify features of a written text that are not readily perceptible by visual inspection. For example, the proportion of each letter appearing in a text, or the frequency of occurrence of certain letter digrams, can be used to identify the language even though the text is not understood. Similarly, the spatial and statistical distribution of optical density in a nucleus can be used to detect chromatin patterns that are not visually perceptible. This notion provides a useful vehicle for achieving advantageous refinements in the detection of pathological change, and of precancerous and cancerous lesions.

[0006] Optical density (OD) of a sample is defined in the art as the logarithm of the ratio of the light incident to the sample and the light transmitted through it. As used in microscopic imagery, optical density is usually expressed in terms of base-ten logarithmic values that range between zero and about 1.80 (because the accuracy of measurement limits near-zero transmission readings). An OD value of zero refers to full transmission, while an OD of 1.80 refers to transmission slightly greater than 1 percent. OD values are conventionally grouped into intervals of 0.10 OD units. For convenience, OD values may be multiplied by a factor of 100, so that computations can be carried out with integers (for example, an OD value of 1.0 is

represented by 100, which corresponds to 10% light transmission).

[0007] As mentioned above, many features may be defined from the statistical and spatial distribution of nuclear chromatin.

5 Global features are computed from the nucleus as a whole. For example, "total optical density" is defined as the sum of all pixel OD values in the nuclear area (i.e., the number of pixels within the outline of a nucleus). This feature is known to be related to the DNA ploidy of the nucleus, a measure of genetic
10 instability and a diagnostic clue for progression toward a pre-malignant or malignant lesion. The variance of optical density within a nucleus is another example of global feature. Other features are local in nature, such as the frequency of occurrence of particular OD values within a certain interval, and have been
15 identified in the art as indicative of tissue condition.

[0008] According to prior-art procedures, the chromatin features characterized using pixel OD values as described above have been reduced to number values representative of a quantitative measure
20 of each feature and of a chromatin or nuclear "signature" representative of a set of features. These numeric values have then been used to provide pathologists with quantitative information available to complement their visual evaluation of tissue slides. For example, as generally described in U.S.

Patent No. 6,204,064, the information derived from the nuclear signature can be used advantageously as a quantitative measure of progression toward a lesion. That is, the physician is provided with information representative of a result formulated by the analytical algorithm built into the diagnostic system (e.g., a numerical value assigned to the nuclear signature calculated by the system and a resulting position in a progression curve). The physician is not provided with an image that has been information-enriched by the system and made available for visual inspection as an aid toward improved diagnostic evaluation. As a result, these prior-art mechanized procedures are often under-utilized by the profession. Moreover, the unique training and ability of pathologists to interpret visual imagery is not employed by the system.

[0009] Accordingly, there is still a need for a mechanized diagnostic system that provides information-enriched images of the tissue of interest in addition to the quantitative information generated by the analytical algorithm. Such a system would enable a pathologist to study histometric and karyometric features that are not visually detectable from the images acquired by the optical system but are rendered visible by the system's enhancing algorithms. Thus, significant chromatin features that might escape visual inspection would become

available for evaluation by the physician.

[0010] Obviously, the usefulness of any optical diagnostic technique is closely tied to the resolution of the imaging system because of the corresponding definition of chromatin patterns and the associated optical-density values. On the other hand, the image resolution of an optical objective is inversely related to its field of view. Therefore, any increase in resolution is accompanied by a corresponding loss of field of view, which greatly affects the system's capability for rapid imaging of useful-size tissue samples. Accordingly, in order to image a tissue sample with the degree of resolution required for cytopathological and histopathological testing, prior-art diagnostic imaging systems have relied on optics with high resolution and a small field of view, and on stitching techniques that yield a compound image from sequentially acquired images of adjacent portions of the tissue sample.

[0011] Typically, the field of view of a suitable microscope (such as a planapochromatic oil-immersion objective with NA 1.40 by Zeiss of Germany) is limited to about 200 microns. Related digital imaging systems enable sampling at six pixels per linear micron (36 pixels per square micron), as shown in Fig. 3, which results in several hundred to a few thousand pixels being

recorded for each nucleus in the sample. Thus, tissue samples (up to about 50x25 mm in size) require multiple imaging steps and subsequent stitching of the resulting partial images of the sample. Because of the difficulty involved in aligning images of adjacent portions of the sample, the resulting composite images of overlapping features are often sufficiently misaligned to become unreliable as a source of optical density information. The stage manipulation and the consequent time required to image an object under high magnification is particularly troublesome in pathology analysis because the diagnostic information in the tissue may be located in only a small portion of the tissue that is being imaged. In addition, the time for image acquisition is large enough to prevent analysis and/or data transmission in real time, which is becoming increasingly important in today's world of mechanization and shared information.

[0012] In view of the foregoing, there also remains a great need for a high-resolution imaging system capable of imaging an entire biological sample in a single scan. This invention is directed at a system that meets this requirement and provides information-enriched diagnostic images for direct visual analysis by practitioners.

BRIEF SUMMARY OF THE INVENTION

[0013] A recent innovation in the field of light microscopy

5 provides a solution to the stitching problem associated with prior-art systems by using a miniaturized microscope array, also referred to herein as a multiple objective array microscope, or simply as an "array microscope." As described in commonly owned International Application PCT/US02/08286, herein incorporated by
10 reference, each miniaturized microscope includes a plurality of optical elements individually positioned with respect to a corresponding image plane and configured to image respective sections of the sample. The array further includes a plurality of image sensors corresponding to respective optical elements and
15 configured to capture image signals from respective portions of the sample.

[0014] In such an array microscope, a linear array of

20 miniaturized microscopes is preferably provided with adjacent fields of view that span across a first dimension of the sample (also referred to herein as y direction), and the sample is translated past the fields of view across a second dimension (x direction) to image the entire sample. Because the diameter of each miniaturized microscope objective is larger than its field

of view (having respective diameters of about 2 mm and 250 μ m, for example), the individual microscopes of the imaging array are staggered in the direction of scanning so that their relatively smaller fields of view are offset over the second dimension but aligned over the first dimension, as illustrated in Fig. 4.

Thus, the detector array provides an effectively continuous linear coverage along the first dimension which eliminates the need for mechanical translation of the microscope in that direction, providing a highly advantageous increase in imaging speed by permitting complete coverage of the sample surface with a single scanning pass along the second dimension.

[0015] Thus, according to the present invention, a biological tissue sample is scanned in a single pass to produce an image of the sample and corresponding optical-density data for each pixel in the image. These data are analyzed in conventional manner using a computerized algorithm to identify, isolate and produce segmented images of nuclei contained in the image of the sample. The OD values corresponding to nuclear chromatin are used to generate and identify numerical patterns known to have statistical significance in relation to the health condition of the biological tissue being tested. These patterns are analyzed to detect, for example, pre-neoplastic changes in histologically normal-appearing tissue that suggest a risk for the development

of a pre-malignant and eventually a malignant lesion. Then, in a novel form of utilization of these numerically derived patterns, this information is converted to a visually perceptible form incorporated into the image of the tissue sample and is displayed for qualitative analysis by a pathologist. Thus, an information-enriched image is produced that displays features previously present but not visually detectible by human observation. In addition, the same or other numerically derived patterns may be used to quantify a property of the sample and produce a quantitative diagnostic result associated with that property.

[0016] Various other purposes and advantages of the invention will become clear from its description in the specification that follows and from the novel features particularly pointed out in the appended claims. Therefore, to the accomplishment of the objectives described above, this invention consists of the features hereinafter illustrated in the drawings, fully described in the detailed description of the preferred embodiment and particularly pointed out in the claims. However, such drawings and description disclose but one of the various ways in which the invention may be practiced.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawings will be provided by the Office upon request and payment of the necessary fee.

[0018] Fig. 1 is an image of a histopathologic section viewed under 600X magnification.

[0019] Fig. 2 is the segmented image of a single nucleus in the section of Fig. 1, wherein the image is further magnified to illustrate chromatin patterns and the effect of pixellation.

[0020] Fig. 3 illustrates the image area corresponding to one square micron of sample surface and typical optical-density numerical data corresponding to that portion of the image.

[0021] Fig. 4 is a simplified schematic representation of the fields of view in an array-microscope layout of 99 miniaturized microscopes in an array of 9 rows of individual microscopes disposed transversely and 11 columns of microscopes disposed at a slight angle, respectively, with respect to the direction of scanning, such that a complete coverage of the sample surface is

achieved during a scan.

[0022] Fig. 5 is a histogram of the relative frequency of occurrence of the vowels in a page of Latin text from Julius Caesar's De Bello Civili.

[0023] Fig. 6 is an example of optical-density values acquired during a microscopic scan of a tissue sample and representing a given statistical pattern, such as the occurrence of certain OD values in a given interval, the occurrence of values in certain OD ranges in adjacent pixels, and the occurrence of a run of three values in the same interval in three adjacent pixels.

[0024] Fig. 7 is the histogram of the relative frequency of occurrence of OD values within each 10-unit interval of the standard 0.0-1.8 OD-unit range for a normal cell line.

[0025] Fig. 8 is the histogram of the relative frequency of occurrence of OD values within each 10-unit interval of the standard 0.0-1.8 OD-unit range for a malignant cell line.

[0026] Fig. 9 illustrates a nuclear signature for normal tissue derived from 93 features of normal ectocervical squamous cells.

[0027] Fig. 10 illustrates a nuclear signature for a lesion derived from 93 features of cells from HSIL (high grade squamous intraepithelial lesion) of the ectocervix.

5 **[0028]** Fig. 11 illustrates the nuclear abnormality assessment of normal cells, shown as the average z-value of all nuclear features measured in normal cells.

10 **[0029]** Fig. 12 illustrates the nuclear abnormality of cells from a lesion, shown as the average z-value of all nuclear features measured in the abnormal cells.

[0030] Fig. 13 illustrates the lesion signature of normal endometrial tissue.

15

[0031] Fig. 14 illustrates the lesion signature of endometrial tissue with atypical hyperplasia.

20 **[0032]** Fig. 15 illustrates the lesion signature of endometrial tissue with carcinoma.

[0033] Fig. 16 shows a bimodal nuclear-abnormality distribution of histologically normal tissue from rectal mucosa from patients with a prior history of colonic adenoma.

[0034] Fig. 17 is a plot showing the percentage of the cases corresponding to the distribution of Fig. 16 having a percentage of nuclei within different groups (0-30%; 31-60%; 61-100%) with nuclear abnormality less than 2.0.

5

[0035] Fig. 18 is a plot of case mean discriminant function scores versus average nuclear abnormality for each case. The cases fall into two groups, with lower abnormality and higher abnormality (on the upper right), which has prognostic relevance and suggests relative risk.

10

[0036] Fig. 19 illustrates the nuclear signature of the cases in the lower risk groups in Fig. 18.

[0037] Fig. 20 illustrates the nuclear signature of the cases in the higher risk groups in Fig. 18.

15

[0038] Fig. 21 is an example showing the position of nuclei in histologically normal-appearing tissue that in fact exhibits pre-neoplastic changes on a progression curve for bladder cancer varying between normal and papillary carcinoma in situ of grade 1 lesions (normal/lesion g_1 , g_2 , and g_3). The symbols g_1 , g_2 and g_3 indicate nuclei measured in histologically normal-appearing tissue from cases of bladder cancer of grades 1, 2 and 3,

20

respectively.

[0039] Fig. 22 illustrates the nuclear signature of the most normal subset of urothelial nuclei (normal/lesion g_1) in Fig. 21).

[0040] Fig. 23 illustrates the nuclear signature of the subset of normal-appearing urothelial nuclei (normal/lesion g_1) showing higher deviations from normal than the subset of Fig. 22.

[0041] Fig. 24 illustrates the nuclear signature of urothelial nuclei from papillary carcinoma in situ of grade 1, showing essentially the same deviation from normal as the subset of Fig. 23.

[0042] Fig. 25 illustrates the image-enhancing property of the invention in an image of a histopathologic section of urothelium taken from a region judged to be histologically normal at some distance from a papillary carcinoma lesion in a bladder. The framed nuclei have signatures indicating pre-neoplastic change.

[0043] Fig. 26 shows a prostate gland duct that appears normal in view of a visible single-layered epithelium and a basal cell layer without gaps.

[0044] Fig. 27 illustrates the nuclear signature of the nuclei imaged in the sample of Fig. 26 (middle) as compared with the signatures from nuclei of normal tissue (above) and from cancerous tissue (below).

5

[0045] Fig. 28 is the sample of Fig. 26 enriched according to the invention by superimposing the signature corresponding to the sample over the image of the epithelium.

10 **[0046]** Fig. 29 is the sample of Fig. 26 enriched according to the invention by digitally staining with the colors yellow and green nuclei found to be abnormal and normal, respectively, by numerical analysis and using a predetermined empirical selection criterion.

15

DESCRIPTION OF THE PREFERRED EMBODIMENTS OF THE INVENTION

[0047] The invention is based on the realization that prior-art advances made in the field of optical histopathology have not
5 been properly utilized to optimize the detection of lesions and other forms of tissue abnormalities. In particular, mechanized prior-art procedures have provided quantitative measurements of tissue abnormalities related of chromatin patterns in the nuclei
10 of cells, but these procedures do not provide this information to the pathologist in a form that takes advantage of his or her highly specialized training in interpreting visual information during the diagnostic process. The numerically derived patterns corresponding to significant chromatin features have been used to provide a quantitative analysis that supplements the
15 pathologist's visual interpretation of the tissue image, but these numerical data have not been used to provide the pathologist additional information for his or her qualitative, visual assessment of the tissue. That is the improvement provided by the present invention.

20
[0048] As used herein, the term "pixel" refers both to the discrete portion of a light detector associated with a signal corresponding to light received thereby and to the corresponding area of the sample from which the light is received. The term

"microscope" is used not only in its generic sense but also with reference to both the array microscope and the individual miniaturized microscopes within the array described in International Application PCT/US02/08286, and it is assumed that the distinction will be apparent to those skilled in the art from the context of the description.

[0049] In order to appreciate the subtlety of the present invention, it is useful to point out the distinction between the concepts of visual-image information and non-visual image information. When one reads a printed text, for example, each letter, word, and the language of the text (if known) are readily recognized. Each conveys visually perceived information. On the other hand, most people could not recognize the frequency of occurrence of a given letter, of a given digram of letters, or of a given sequence of letters. Yet, each of these conveys statistically significant non-visual information that can be utilized to characterize features of the text. For instance, vowels are known to occur with different frequencies of occurrence in different languages. Therefore, a histogram of the relative frequency of occurrence of each vowel, as illustrated in Fig. 5, could be used by a linguist to identify the language or, for example, to place it within an evolutionary scale of related languages. Thus, it is clear that even information that is not

readily perceived by visual inspection (referred to herein as "non-visual" information) can be converted into visually perceived information, such as the histogram of Fig. 5, which provides additional knowledge that is significant and can be used advantageously by experts in the field of linguistics. For instance, if the vowel "a" in a given text occurs at the frequency expected for a certain language, it may be printed in green. If it occurs at an increased frequency, say twice the expected rate, it may be printed in red.

[0050] Another example lies in the field of music. When played, music provides audibly perceptible patterns as well as hidden statistical dependencies in the sound pattern that escape recognition even by the most sophisticated listener. Yet these undetected sound patterns may carry information which, if identified in the form of statistically significant data, could greatly enhance the process of studying and analyzing musical structure. For instance, non-audibly detectible patterns may be identified in terms of numerical values based on signals contained in a digital recording, isolated from the recording, and, for evaluation by an analyst, they may be converted into music notes, if the analyst is particularly skilled at perceiving audible information. They could also be converted into color in the printed musical scores. This information could be used by

musicologists to study the original musical composition to further refine their understanding of its structure. For example, such a process could be used to aid in the identification of the author of a newly discovered score by
5 comparing numerically derived patterns identified in the score with the same patterns found in the music of known authors. Thus, again, information that is not directly detectable by an expert can be identified by a machine using numerically derived, statistically significant patterns. The same information can
10 then be converted to a detectable form which an expert in the field can analyze and take into account in a qualitative analysis of the original work.

[0051] With reference to histopathology, Fig. 6 provides an
15 example of how optical-density values may be used according to the invention. Each line 11-13 represents a pixel sequence of OD values (scaled by 100) acquired during a microscopic scan of a tissue sample. Each value corresponds to a pixel within the area of a nucleus in the image and reflects the optical density
20 imparted by the chromatin in that pixel. The first line 11 illustrates the occurrence of scaled OD values in the 30-39 interval; the second line 12 illustrates the occurrence of values in the 40-49 and 70-79 ranges in adjacent pixels; and the third line 13 illustrates the occurrence of a run of three values in

the 40-49 range in three adjacent pixels. Applied to the entire nuclear region, each of these variables provides a useful analytical tool for characterizing the state of a cell. For example, Figs. 7 and 8 show the difference between the histograms of the relative frequency of occurrence of OD values within each 10-unit interval for normal and malignant cell lines, respectively. The figures show that the optical density of malignant cells is skewed toward higher values in a statistically significant manner that reflects the existence of dense chromatin granules associated with malignancy. Thus, this is an OD feature that provides useful patterns to identify and quantify the progression of a lesion, as disclosed in U.S. Patent No. 6,204,064.

[0052] Many other OD features have been used in the art to identify useful patterns for diagnostic applications. For example, other first-order statistical variables may consist of how often consecutive OD values in a line of pixels fall within a certain interval. Higher-order statistical variables may similarly be used, such as the slopes of OD values between adjacent pixels. Finally, additional specific features have been utilized to summarize statistical properties, such as run-length uniformity or non-uniformity, OD clumpiness, shape of the OD histogram, and other descriptive features. These features have

been used in the art in various manners in order to produce a nuclear signature which, when compared to the nuclear signature of a normal tissue, provides a quantitative assessment of the condition of the tissue. See, for example, P.H. Bartels et al.,
5 "Extraction and evaluation of information from digitized cell images," in *Mammalian cells: probes and problems*, Richmond et al. ed., Oak Ridge, Tenn., Technical Information Center, 1975:15-28; T. Young et al., "Characterization of chromatin distributions in cell nuclei," *Cytometry* 1986, 7:467-474; E. Bengtsson et al.,
10 "Densitometry, Morphometry and Texture Analysis as Tools in Quantitative Cytometry and Automated Cancer Screening," Grohs and Husain OAN, editors, 1994. p. 21-43; and A. Doudkine et al., "Nuclear texture measurements in image cytometry," *Pathologica* 1995, 87:286-299.

15
[0053] According to the present invention, any number of these same features is used instead to provide a visually perceptible representation of otherwise undetectable information. Because of its visible form, this information is made available to the
20 pathologist for a qualitative assessment of the tissue to be used in addition to the images provided in normal practice. For example, the sum of the optical density within the nucleus of the cells in a given tissue sample may be measured as a variable indicative of the presence or absence of a lesion based on a

statistical comparison with normal and malignant cells. This information may then be used to provide an information-enriched image by adding a color to each nucleus indicative of its true condition based on the statistical analysis.

5

[0054] In practice, a plurality of OD features is normally selected for analysis and is ordered in an arbitrary (but consistently maintained to allow comparison) predetermined sequence. To form a nuclear signature, a set of reference nuclei
10 from normal tissue of a given organ site is used. The mean and the standard deviation in the reference cells are computed for each of the plurality of selected features. For any nucleus in the data set to be analyzed, the absolute difference to the mean in the reference data set is computed for each feature and is
15 "normalized" by dividing it by the standard deviation found in the reference data set for that feature. In other words, the feature values are expressed in terms of so-called z-values (that is, in units of standard deviation, also referred to in the art as z-scores). Thus, the reference data set will produce a z-
20 value for each feature that depends on the shape of the probability function for that feature. Fig. 9 illustrates such a nuclear signature for normal tissue derived from 93 features of normal squamous-cell tissue. For a Gaussian, symmetrically distributed feature, the z-value will be 0.67; that is, a little

less than one standard deviation and a value for which 50% of all nuclei have a lower, and 50% a higher, deviation. For skewed distributions, it may be higher or lower, as seen for some features in Fig. 9.

5

[0055] Repeating the same statistical exercise with the same 93 OD features using data obtained from nuclei from a known lesion, a clearly distinct profile is obtained, as shown in Fig. 10, which illustrates the usefulness of these features in assessing the pathologic condition of cells. In order to reduce the dimensionality of the nuclear-signature representation, an average of the z-values over all features may be used instead of the entire profile. Thus, a single number for each nucleus is obtained that describes the average deviation of the nuclear features from normal, referred to herein as the "nuclear abnormality." Figs. 11 and 12 illustrate the results of such an averaging process. If a histogram is constructed to reflect the nuclear abnormality of each measured nucleus in a sample, the resulting distribution is referred to as a "lesion signature."

20

[0056] Fig. 13 illustrates the lesion signature of normal endometrial tissue. Most cells exhibit a nuclear abnormality in the lower, normal value range (see Fig. 11); very few cells show a nuclear-abnormality value above 1.0; and none have a value

above 2.0. In contrast, tissues with lesions show progressively different lesion signatures with more cells with high nuclear abnormality and fewer cells with normal nuclear abnormality.

Fig. 14 illustrates the lesion signature of a tissue sample with atypical hyperplasia, showing a significant incidence of nuclear abnormality values above 2.0 and a reduced number of normal values. Fig. 15 illustrates the lesion signature of a tissue with carcinoma, wherein the nuclear abnormality is seen spread through a wide range of higher values.

[0057] An important aspect of the invention lies in the recognition that these differences in OD features appear not only in cells that produce visually recognizable image features (that is, features that the average pathologist identifies visually as significant for a diagnosis), but they appear also in apparently healthy tissue. Accordingly, the invention is particularly useful in the detection of such pre-neoplastic changes in histologically normal-appearing tissue that suggest a risk for the development of a pre-malignant and a potentially malignant lesion. It has also been shown in some organ sites that the degree of pre-neoplastic change corresponds to the grade (an established measure of progression) of existing lesions at different locations in the same organ. Thus, the detection of pre-malignant change in a biopsy taken from tissue that appears

histologically normal to the human eye may point to the presence of an existing pre-malignant or malignant lesion in the same organ, or it may suggest a risk for the recurrence of a lesion removed by an earlier intervention. It has also been shown that nuclear-abnormality values in breast lesions are highly correlated with risk values accepted by pathologists for the development of infiltrating disease.

[0058] Thus, in general and according to the invention, a decision rule may be derived from cases of known outcome and then used to assign a risk value to new cases. For example, a data set of normal tissue for a given organ of interest is established and then nuclear signatures for a large number of nuclei from cases of known outcome are determined, as explained above. The distribution of nuclear-abnormality values is then computed for the entire set of nuclei. For instance, Fig. 16 shows the nuclear-abnormality distribution of histologically normal rectal mucosa from patients with a prior history of colonic adenoma (a known potential precursor lesion for colonic adenocarcinoma). It is evident from the figure that the nuclei fall into two distinct modes, one with moderate deviation and the other with marked deviation from normal. Thus, as illustrated in the figure, an arbitrary but judiciously selected threshold T may be used to separate the two modes at a deviation from normal, for example,

equal to two feature-averaged standard deviations.

[0059] Fig. 16 clearly shows that the images of cells that appear normal to visual inspection carry visually undetectable

5 information that can be used according to the invention to reveal otherwise undetected abnormalities. To that end, the data

corresponding to Fig. 16 are further utilized to calculate the proportion of nuclei in each case that fall into one or the other deviation mode (i.e., below or above the threshold of 2.0). Fig.

10 17 illustrates such a plot with reference to the cases used for Fig. 16. As would be clear to one skilled in the art, the cases with the highest proportion of nuclei in the group below the threshold represent the most normal cases. On the other hand,

the cases with a small percentage of nuclei below the threshold

15 represent cases with greater deviation from normal. Fig. 17, for example, shows 51 percent of the reference cases having at least 67 percent of their nuclei with moderate deviation from normal

(i.e., less than 2.0 nuclear abnormality), illustrated by area A;

33 percent of the cases have only 0-30 percent of their nuclei in the same range (area B); and the remaining 16 percent of cases have between 30 and 67 percent of their nuclei in the moderate deviation range (area C).

[0060] The statistical significance of this difference may be

further refined by applying a nonparametric test to the two subsets of data at each end of the nuclear abnormality plot. For example, a Kruskal Wallis test may be conducted in conventional manner to determine which of the OD features (93 were used in the example) are statistically different between the two data subsets at a certain significance level, and to distinguish cases with moderate and high deviation from normal. For instance, features from different feature groups may be selected from among the features testing at a p-level of 0.005 or less. Using a standard stepwise discriminant algorithm, the mean discriminant function score is computed for each case over all of its nuclei. These scores are then plotted as a function of the corresponding average nuclear abnormality for each case, as illustrated in the plot of Fig. 18, which also shows the 95-percent confidence-level ellipses for the bivariate mean vectors and the 90-percent tolerance ellipses for case mean values.

[0061] The location of a data point representing a case in the plot of Fig. 18 indicates the degree of pre-neoplastic change in the tissue. Similarly, Figs. 19 and 20 show the distinctly different nuclear signatures of the cases in the lower and higher risk groups in Fig. 18. Thus, the grade of a coexisting pre-malignant or malignant lesion and/or the associated risk of possible recurrence of a pre-malignant lesion may be assessed by

the procedure. It is noted that the deviation from normal in cases with pre-neoplastic change is not necessarily expressed in every nucleus. Further processing of the data sets with a non-supervised learning algorithm (P-index) may show that the data set with an average deviation between normal and a low grade premalignant or malignant lesion actually consists of two subsets. For example, one may be formed by nuclei almost indistinguishable from normal, whereas the other is practically indistinguishable from low grade malignant nuclei. When present in only 25 to 50 percent of the nuclei in seemingly histologically normal tissue, these deviant cells are not detected by visual assessment but nonetheless exhibit a nuclear abnormality consistent with pre-malignant and malignant lesions, as these are quantified in the art. Fig. 21 illustrates an example of this phenomenon by showing the position of nuclei with pre-neoplastic change on a progression curve for bladder cancer varying between "normal" and papillary carcinoma in situ (CIS) of grade 1 lesions. This figure shows that the statistically derived numerical measure of nuclear abnormality used to practice the invention tracks the nuclear-abnormality of cells graded as diseased as a result of visual diagnosis by a pathologist.

[0062] Figs. 22-24 show the heterogeneity of the nuclei measured in the histologically normal-appearing urothelium of Fig. 21.

The nuclear signature of Fig. 22 is that of the most normal subset of urothelial nuclei in the cases falling under the label normal/lesion in Fig. 21. (The notation "normal/lesion" is used in the art to label nuclei measured in histologically normal-
5 appearing tissue in an organ harboring a premalignant or a malignant lesion at a different location.) Fig. 23 is the nuclear signature of the about 25% of nuclei showing higher deviations from normal in the normal/lesion data set. This figure shows that the deviations are virtually indistinguishable
10 from those seen in nuclei from grade 1 papillary carcinoma (labeled CA), shown in Fig. 24.

[0063] Thus, the invention provides a clearly significant measure of visually undetectable nuclear features that may be extracted
15 from the optical-density information contained in the microscope image of a tissue sample and used to provide an information-enriched image for the pathologist's qualitative assessment of the tissue. This can be done, for example, by associating a given marker to each cell in the image found to correspond to a
20 statistically significant measure of abnormality based on optical-density features. Different colors or computer graphic symbols (at continuous or discrete grading), for instance, may be assigned to all cells falling within a certain range of such measure and the resulting image may be displayed to give the

pathologist a picture of which normal-appearing nuclei in fact show signs of statistically significant degradation and the degree of such degradation (as represented by differences in assigned colors or computer-graphic means). The pathologist then, based on his or her experience, may use this new picture of the tissue as further information for a qualitative assessment of the tissue's condition. The new picture with such markers may also be combined with the original image to provide an information-enriched image for diagnostic use by the pathologist. The invention is best illustrated by the following examples.

[0064]

Example 1

Fig. 25 shows a histopathologic section of urothelium taken from a region judged to be histologically normal at some distance from a papillary carcinoma lesion in a bladder. Nine nuclei in the image (identified by highlighted boundaries in the figure) were selected as suitable for analysis for their in-focus position. Upon digital analysis, two of these nuclei showed a nuclear signature resembling that of papillary carcinoma, even though visually the section was normal. Utilizing a computer graphic marker in the form of a outlining box, the diagnostic information is revealed directly on the image, thereby providing a visual, quantitative diagnostic clue not otherwise detectable from the image.

[0065]

Example 2

A tissue sample from a prostate gland duct was sectioned, fixed and stained with hematoxylin/eosin. The resulting slide was scanned with a high numerical aperture microscope objective to produce an image of the sample, as shown in Fig. 26. The image exhibits all criteria for a normal gland, such as a single-layered epithelium and a basal cell layer without gaps. However, a numeric analysis of the chromatin pattern of the nuclei revealed an overall signature known to be typical of a "normal appearing gland" in the vicinity of a prostatic intraepithelial neoplasia (PIN) lesion. Fig. 27 illustrates the signature of the imaged sample (seen in the middle of the figure) in comparison with the signatures of normal tissue (above) and cancerous tissue (below).

[0066] According to one approach, the signature profile of the epithelium is displayed superimposed on the image of the stained sample to reveal the diagnostic clue that is not otherwise visually perceptible, as illustrated in Fig. 28. According to another, preferred approach, individual nuclei within the epithelium are digitally stained according to the signature profile computed for each nucleus using different colors to separate abnormal nuclei from normal ones according to a predetermined selection criterion. Fig. 29 illustrates this

approach where nuclei found to be abnormal by numerical analysis are digitally stained yellow and normal nuclei are stained green.

5 **[0067]** A comparison of the images produced by the invention with the original images for the same tissue illustrate the additional information available to assist the diagnostic process by a pathologist. While the marked cells are not intended to provide a quantitative result indicative of the tissue condition, the simple existence of the abnormalities reflected by them and the
10 spatial distribution of such abnormalities will provide an additional qualitative tool for the practitioner's consideration while making a diagnostic assessment of the tissue. Thus, the nature of the information produced by the present invention is not to be confused with the quantitative assessments described in
15 the prior art.

[0068] Various other changes in the details, steps and components that have been described may be made by those skilled in the art within the principles and scope of the invention herein
20 illustrated and defined in the appended claims. For example, the invention has been described in the context of pathology practice with a conventional transillumination system, wherein a tissue sample is fixed to a slide for staining and microscopic examination by a pathologist. The image normally seen by the

pathologist is enhanced as a result of numerical analysis in order to render visible otherwise visually imperceptible information. Clearly, the same inventive concepts can be used with other image-producing techniques, whether based on transmitted, absorbed, or emitted light (such as fluorescence-emission imaging, natural or induced. Similarly, the invention could be carried out with a combination of diagnostic clues obtained from numerical analysis of different kinds of imaging techniques.

[0069] Therefore, while the present invention has been shown and described herein in what is believed to be the most practical and preferred embodiments, it is recognized that departures can be made therefrom within the scope of the invention, which is not to be limited to the details disclosed herein but is to be accorded the full scope of the claims so as to embrace any and all equivalent processes and products.